Room 14-0551
77 Massachusetts Avenue
Cambridge, MA 02139
Ph: 617.253.5668 Fax: 617.253.1690
Email: docs@mit.edu
http://libraries.mit.edu/docs

September, 1981 LIDS-P-1141

HIERARCHICAL AGGREGATION OF LINEAR SYSTEMS WITH MULTIPLE TIME SCALES

M. Coderch,* A.S. Willsky,* S.S. Sastry,*
and D.A. Castanon*

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

0. INTRODUCTION

Models of large scale systems typically include weak couplings between some states. This in turn leads to the evolution of different portions of the system at different time scales. Intuition suggests that the analysis of phenomena at one time scale is made tractable (simplified) by assuming constancy of variables at slower time scales and steady state values for variables at faster time scales. It is this intuition of a hierarchy of approximations that we make precise in this paper. Mathematical models of interconnected power systems have variations on several time scales -nearly instantaneous adjustment of (PV,PQ) load bus angles and voltages, dynamics of the swing equations, voltage regulator and generation variation dynamics, generation set point changes are examples of progressively slower dynamics.

The presence of uncertainties, load fluctuations, rare catastrophic events in the power system clearly necessitates stochastic models for large scale power systems. Model simplification then consists of identifying all the time scales present in the given model and presenting approximate models valid (uniformly over time, in a sense made precise in the paper) at each time scales.

The class of models considered in this paper is of the linear time invariant form (1-1). Study of this (deterministic) equation is relevant in the study of hierarchical aggregation of finite state Markov processes with weak couplings (symptomatic of multiple time scales), described by small parameter ε. Details of the applications of our results to this context will be presented in [3].

1. PROBLEM FORMULATION

We consider here linear time-invariant systems of the form

$$\dot{x}^{\varepsilon} = A_0(\varepsilon)x^{\varepsilon} \qquad x(0)=x_0 \tag{1.1}$$

where $x^{\varepsilon} \in R^n$ and the matrix $A_0(\varepsilon)$ is an analytic function of ε*, of normal rank d except at $\varepsilon=0$

$$A_0(\varepsilon) = \sum_{k=0}^{\infty} A_{0k}\varepsilon^k \tag{1.2}$$

we analyse the asymptotic behavior of $x^{\varepsilon}(t)$ as $\varepsilon \downarrow 0$ on the time interval $[0,\infty[$. In general

$$\lim_{\varepsilon\downarrow 0}\ \sup_{t\geq 0} ||x^{\varepsilon}(t)-x^{o}(t)||\neq 0$$

so that asymptotic behavior at several time scales needs to be considered - $x^{\varepsilon}(t)$ is said to have well defined behavior at time scale $t/g(\varepsilon)$ (where $g(\varepsilon)$ is a monotone increasing C^o function on $[0,\varepsilon_o]$ with $g(0)=0$) if there exists a bounded continuous function $y_k(t)$, called the evolution at that time scale, such that

$$\lim_{\varepsilon\downarrow 0}\ \sup_{\delta<t\leq T} ||x^{\varepsilon}(t/g(\varepsilon))-y_k(t)||=0 \qquad \delta>0,\ T<\infty \tag{1.3}$$

In this paper we give tight sufficient conditions under which the multiple time scale behavior of $x^{\varepsilon}(t)$ can be fully described by its evolutions at time scales t/ε^k for integers $k=0,1,\ldots,m$. These evolutions are used to:

(i) provide a set of reduced order models valid at different time scales.

(ii) provide an asymptotic approximation to $x^{\varepsilon}(t)$ valid uniformly on $[0,\infty[$.

2. NOTATION

Given $A \in R^{nxn}$, $\mathcal{R}(A)$ and $\mathcal{N}(A)$ denote the range and null space of A. $\rho(A)$ denotes the resolvent set of A, i.e. the set of $\lambda \in C$ such that the resolvent, denoted $R(\lambda,A):=(A-\lambda I)^{-1}$, is well defined.

If $\lambda=0$ is an eigenvalue of algebraic multiplicity m, then the Laurent series of $R(\lambda,A)$ at

---

* Research supported in part by the DOE under grant ET-76-C-01-2295. The authors are thankful to B.Levy and J.C. Willems for the helpful discussions.

* Formally, our results hold even if $A_0(\varepsilon)$ is only assumed to have an asymptotic series (see, for eg. [4]) of the form (1.2), provided that $A_0(\varepsilon)$ has constant rank d for $\varepsilon\in]0,\varepsilon_0]$.

$\lambda=0$ has the form (see [5])

$$R(\lambda,A)=\frac{-P(A)}{\lambda}-\sum_{k=1}^{m-1}\lambda^{-k}D(A)^k+\sum_{k=0}^{\infty}\lambda^k S(A)^{k+1} \qquad (2.1)$$

where $P(A)$, the <u>eigenprojection</u>; $D(A)$, the <u>eigen-nilpotent</u> and $S(A)$ are defined by

$$P(A):=\frac{-1}{2\pi i}\int_\gamma R(\lambda,A)\,d\lambda \qquad (2.2)$$

$$D(A):=\frac{-1}{2\pi i}\int_\gamma \lambda R(\lambda,A)\,d\lambda \qquad (2.3)$$

$$S(A):=\frac{-1}{2\pi i}\int_\gamma \lambda^{-1}R(\lambda,A)\,d\lambda \qquad (2.4)$$

with γ a positively oriented closed contour enclosing 0 but no other eigenvalue of A.

A is said to have <u>semi-simple null structure</u> (SNS) if $D(A)=0$. In that case, $R^n = \mathcal{R}(A) \oplus \mathcal{N}(A)$ and $P(A)$ is the projection onto $\mathcal{N}(A)$ along $\mathcal{R}(A)$. A is said to be <u>semi-stable</u> if it has SNS and all its non-zero eigenvalues are in the open left half plane.

If A is semistable, then $\lim_{t\to\infty} e^{At}=P(A)$ and further

$$S(A)=-\int_0^\infty (e^{At}-P(A))\,dt = (A+P(A))^{-1}-P(A).$$

Since $S(A)$ is a generalized inverse of A ($S(A)Ax=x=AS(A)x$ for all $x\in\mathcal{R}(A)$) we denote it $A^\dagger$.

3. STATEMENT OF MAIN RESULT

We present here a uniform (in t) asymptotic approximation of $e^{A_0(\varepsilon)t}$ involving behavior at time scales t/ε^k ; $k=0,1,\ldots,m$. The proof relies on results in [5] and is outlined in Section 4.

For our development, we need an array of matrices A_{ik}, $i\geq 0$, $k\geq 0$ starting from the A_{0k} of (1.2), constructed recursively ((k+1)th row from kth row) by the following formula

$$A_{k+1,\ell}=-\sum_{p=1}^{\ell+1}(-1)^p \sum_{\substack{\nu_1+\ldots+\nu_p=\ell+1\\ k_1+\ldots+k_{p+1}=p-1\\ \nu_i\geq 1,\ k_i\geq 0}} S_k^{(k_1)}A_{k,\nu_1}S_k^{(k_2)}\cdots A_{k,\nu_p}S_k^{(k_{p+1})}$$

where $S_k^{(0)}=-P(A_{k,0})$ and $S_k^{(k_i)}=(A_{k,0}^\dagger)^{(k_i)}$.

<u>Remarks</u>: (i) The computation of $A_{k+1,\ell}$ requires only $A_{k,0}, A_{k,1},\ldots,A_{k,\ell+1}$ so that it proceeds triangularly as shown in Table 1.

(ii) of special interest to us in the sequel is the structure of $A_{00}, A_{10}, A_{20},\ldots$, since they determine the leading term in the asymptotic expansion of $e^{A_0(\varepsilon)t}$. For these matrices, (3.1) can be simplified considerably. (see Remark (ii) after the Theorem and Corollary).

<u>Theorem</u>

Let $A_0(\varepsilon)$ be a matrix with SNS of the form (1.2) of normal rank d except at $\varepsilon=0$. If A_{00}, A_{10}, A_{m0} are semistable with rank A_{00}+ rank $A_{10}+\ldots+$ rank $A_{m0}=d$ then

$$R^n = \mathcal{R}(A_{00}) \oplus \ldots \oplus \mathcal{R}(A_{m0}) \oplus \mathcal{N} \qquad (3.2)$$

where $\mathcal{N}=\bigcap_{k=0}^{m}\mathcal{N}(A_{k0})$.

Further let P_k, for $k=0,1,\ldots,m$, be the projection onto $\mathcal{N}(A_{k0})$ defined by (2.2) and $Q_k = I-P_k$.

Then

$$\lim_{\varepsilon\downarrow 0}\ \sup_{t\geq 0}\left\|e^{A_0(\varepsilon)t}-\phi(\varepsilon,t)\right\|=0 \qquad (3.3)$$

where $\phi(\varepsilon,t)$ is any one of the four expressions below

$$\sum_{k=0}^{m} e^{A_{k0}\varepsilon^k t}Q_k + P_0P_1\cdots P_m \qquad (3.4)$$

$$\sum_{k=0}^{m} e^{A_{k0}\varepsilon^k t} - mI \qquad (3.5)$$

$$\prod_{k=0}^{m} e^{A_{k0}\varepsilon^k t} \qquad (3.6)$$

$$e^{\sum_{k=0}^{m} A_{k0}\varepsilon^k t} \qquad (3.7)$$

With this theorem in hand, the entire multiple time-scale structure of (1.1) can be read off as follows:

Corollary

Under the conditions of Theorem 1, $x^{\varepsilon}(t)$ of (1.1) has the following multiple-time scale properties:

$$\text{(i)}\ \lim_{\varepsilon \downarrow 0} \sup_{0<\delta\leq t\leq T<\infty} \left|\left|x^{\varepsilon}(t/\varepsilon^{k}) - \pi_k\, e^{A_{k0}t} x_0\right|\right| = 0 \qquad (3.8)$$

for $\delta>0$, $T<\infty$ and $k=0,1,\ldots,m-1$

$$\text{(ii)}\ \lim_{\varepsilon \downarrow 0} \sup_{0<\delta\leq t<\infty} \left|\left|x^{\varepsilon}(t/\varepsilon^{m}) - \pi_m\, e^{A_{m0}t} x_0\right|\right| = 0 \qquad (3.9)$$

for $\delta>0$.

where $\pi_0=I$ and $\pi_k=P_0P_1\ldots P_{k-1}$ for $k=1,\ldots,m$.

Equation (3.8) implies the results of [1] and [2] where the authors analysed the convergence $e^{A_0(\varepsilon)t/\varepsilon^{s}}$ for <u>fixed</u> s and over <u>compact</u> time intervals of the form $[0,T]$.

<u>Remarks</u>: (i) The requirement of semi-stability of the matrices $A_{00}, A_{10}, \ldots, A_{m0}$ to obtain well defined behavior at time scales t/ε^{k} is a tight sufficient condition. Examples showing the failure of the theorem without these assumptions, are given in Section 5.

(ii) It is important to be able to calculate the A_{k0} for $k=0,1,\ldots,m$ from the given data A_{00}, A_{01} $A_{02}, \ldots$ of (1.?), without having to obtain the complete matrix of Table 1. This can be done by a variety of methods. One approach that is successful is the formal asymptotics of [7] relating the A_{k0} to Toeplitz matrices constructed with the $\{A_{0i}\}_{i=0}^{\infty}$. Connection is made therein with the Smith McMillan zero of $A_0(\varepsilon)$ at $\varepsilon=0$. In particular <u>m is proven to be the order of the Smith McMillan zero of $A_0(\varepsilon)$ at $\varepsilon=0$</u>. The construction of the A_{k0} from the A_{0i} proceeds as follows:

A_{10} is the null extension to R^n of A_{01} mod $R(A_{00})/N(A_{00})$, i.e, $A_{10}=P_0A_{01}P_0$, where P_0 is defined in the theorem. Pictorially A_{10} is the null extension to R^n of $\hat{A}_{10}$ obtained as below

$$\begin{tikzcd}
N(A_{00}) \arrow[r, hook, "i_0"] \arrow[drr, dashed, "\hat{A}_{10}"'] & R^n \arrow[r, "A_{01}"] & R^n \arrow[d, "P_0"] \\
 & & R^n|R(A_{00}) \cong N(A_{00})
\end{tikzcd}$$

(here, i_0 is inclusion)

A_{20} is the null extension to R^n of $A_{02}-A_{01}A_{00}^{\dagger}A_{01}$ mod $R(A_{00})$ mod $R(\hat{A}_{10})|N(\hat{A}_{10})$, i.e., A_{20} is given by

$$P_1P_0(A_{02}-A_{01}A_{00}^{\dagger}A_{01})P_0P_1:$$

where P_1 is defined as in the theorem Pictorially, A_{20} is the null extension to R^n of $\hat{A}_{20}$,

$$\begin{tikzcd}
N(\hat{A}_{10}) \arrow[r, hook, "i_1"] \arrow[ddrrr, dashed, "\hat{A}_{20}"'] & (A_{00}) \arrow[r, hook, "i_0"] & R^n \arrow[r, "A_{02}-A_{01}A_{00}^{\dagger}A_{01}"] & R^n \arrow[d, "P_0"] \\
 & & & N(A_{00}) \cong R^n|R(A_{00}) \arrow[d, "P_1"] \\
 & & & N(\hat{A}_{10}) \cong N(A_{00})|R(\hat{A}_{10})
\end{tikzcd}$$

and so on. The reader may refer to [7] for complete proof and details as well as the connections between $A_{10}, A_{20}, A_{30}, \ldots$ and the Toeplitz matrices

$$\begin{bmatrix} A_{01} & A_{00} \\ A_{00} & 0 \end{bmatrix}, \quad \begin{bmatrix} A_{20} & A_{10} & A_{00} \\ A_{01} & A_{00} & 0 \\ A_{00} & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} A_{03} & A_{02} & A_{01} & A_{00} \\ A_{02} & A_{01} & A_{00} & 0 \\ A_{01} & A_{00} & 0 & 0 \\ A_{00} & 0 & 0 & 0 \end{bmatrix}$$

and so on.

(iii) The reader should observe using the formulae in remark (ii) above that even if $A_0(\varepsilon)=A_{00}+\varepsilon A_{01}$ the system can exhibit time scales of order t/ε^2, t/ε^3 and so on. This is not a widely appreciated fact.

(iv) <u>Reduced order models</u> It follows from (3.8) and (3.9) that the evolution of $x^{\varepsilon}(t)$ at time scales t/ε^{k}, $k=0,1,\ldots,m$ is given by

$$y_k(t) = e^{A_{k0}t}\pi_k x_0 \qquad k=0,1,\ldots,m$$

Thus, $x^{\varepsilon}(t)$ may be represented asymptotically by the following expression uniformly valid for $t\geq 0$.

$$x^{\varepsilon}(t) = \sum_{k=0}^{m} y_k(\varepsilon^{k}t) + (I - \sum_{k=0}^{m} \pi_k) + o(1)$$

From the direct sum decomposition (3.2) of the theorem, it is clear that a basis $T \in R^{n\times n}$ can be chosen such that

$$e^{A_0(\varepsilon)t} = T\begin{bmatrix} I & & & & \phi \\ & e^{\tilde{A}_m \varepsilon^m t} & & & \\ & & \ddots & & \\ & & & e^{\tilde{A}_1 \varepsilon t} & \\ \phi & & & & e^{\tilde{A}_0 t} \end{bmatrix} T^{-1} + o(1) \qquad (3.10)$$

where $\tilde{A}_0,\ldots,\tilde{A}_m$ are full rank square matrices representing the non zero portions of $A_{00},\ldots,A_{m0}$ in the new basis. (3.10) shows that the system (1.1) decouples asymptotically into a set of subsystems evolving at different time scales governed by the reduced order dynamics of $\{A_k\}_{k=0}^{m}$.

(v) Two time scale systems have been the focus of considerable effort by Kokotovic and coworkers (see [6], for example). It is easy to check in our framework that the assumptions in their systems guarantee precisely two time scales.

(vi) The significance of each row of matrices in Table 1 is discussed in the comment following the proof of the theorem in the next section.

4. Sketch of the proof

The proof relies heavily on results in [5] which are summarized in the following lemma:

Lemma

Let $A_0(\varepsilon)$ have the expression (1.2).

(i) If $\lambda \in \rho(A_{00})$ then for ε small enough $\lambda \in \rho(A_0(\varepsilon))$ and $R(\lambda, A_0(\varepsilon))$ satisfies

$$R(\lambda, A_0(\varepsilon)) = R(\lambda, A_{00}) + \sum_{k=0}^{\infty} \varepsilon^k R^{(k)}(\lambda) \qquad (4.1)$$

for some $\{R^k(\lambda)\}_{k=0}^{\infty}$. Further, the convergence of $R(\lambda, A_0(\varepsilon))$ to $R(\lambda, A_{00})$ as $\varepsilon \downarrow 0$ is uniform on compact subsets of $\rho(A_{00})$.

(ii) Let γ be a closed contour enclosing only the zero eigenvalue of A_{00}. The matrix

$$P_0(\varepsilon) = \frac{-1}{2\pi i} \int_{\gamma} R(\lambda, A_0(\varepsilon))\, d\lambda \qquad (4.2)$$

is well defined for ε sufficiently small and equals the sum of the eigenprojections of all the eigenvalues of $A_0(\varepsilon)$ that converge to 0 (the zero-group) as $\varepsilon \downarrow 0$. $P_0(\varepsilon)$ is analytic in ϵ and commutes with $A_0(\varepsilon)$. Further, $\mathcal{R}(P_0(\varepsilon))$ and $\mathcal{R}(P_0)$ are isomorphic.

(iii) If A_{00} has SNS then $A_1(\varepsilon)$ defined by

$$A_1(\varepsilon) := \frac{A_0(\varepsilon)P_0(\varepsilon)}{\varepsilon} \qquad (4.3)$$

has expansion

$$A_1(\varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k A_{1,k} \qquad (4.4)$$

where the $A_{1,k}$ are obtained from the A_{0k} by (3.1).

Proof of theorem

Define $Q_0(\varepsilon) = I - P_0(\varepsilon)$. Then, note that

$$e^{A_0(\varepsilon)t} = P_0(\varepsilon)e^{A_1(\varepsilon)\varepsilon t} + Q_0(\varepsilon)e^{A_0(\varepsilon)t} \qquad (4.5)$$

By the assumption that A_{00} has SNS, $A_1(\varepsilon)$ has the expansion (4.4). Repeating the manipulation required to yield (4.5) for its first term and using the SNS of A_{10}, we obtain

$$e^{A_0(\varepsilon)t} = P_0(\varepsilon)P_1(\varepsilon)e^{A_2(\varepsilon)\varepsilon^2 t} + Q_1(\varepsilon)e^{A_1(\varepsilon)\varepsilon t} + Q_0(\varepsilon)e^{A_0(\varepsilon)t} \qquad (4.6)$$

where $P_1(\varepsilon)$ is the total-projection on the zero group of $A_1(\varepsilon)$, $Q_1(\varepsilon) = I - P_1(\varepsilon)$ and

$$A_2(\varepsilon) = \frac{P_1(\varepsilon)P_0(\varepsilon)A_0(\varepsilon)}{\varepsilon^2}.$$

Note that

$$\mathcal{R}(Q_0(\varepsilon)) \oplus \mathcal{R}(Q_1(\varepsilon)) \oplus \mathcal{R}(P_0(\varepsilon)P_1(\varepsilon)) = \mathbb{R}^n$$

Under the SNS assumption on $A_{10}, A_{20}, \ldots, A_{m0}$ this procedure may be repeated m times to yield

$$e^{A_0(\varepsilon)t} = P_0(\varepsilon)P_1(\varepsilon)\ldots P_m(\varepsilon)e^{A_{m+1}(\varepsilon)\varepsilon^{m+1}t} + \sum_{k=0}^{m} Q_k(\varepsilon)e^{A_k(\varepsilon)\varepsilon^k t} \qquad (4.7)$$

$$\mathbb{R}^n = \mathcal{R}(Q_0(\varepsilon)) \oplus \mathcal{R}(Q_1(\varepsilon)) \oplus \ldots \oplus \mathcal{R}(Q_m(\varepsilon)) \oplus \mathcal{R}(P_0(\varepsilon)\ldots P_m(\varepsilon)) \qquad (4.8)$$

By the part (ii) of the lemma above

$$\dim \mathcal{R}(Q_k(\varepsilon)) = \operatorname{rank} A_{k0}.$$

further, since

$$\operatorname{rank} A_{00} + \operatorname{rank} A_{10} + \ldots + \operatorname{rank} A_{m0} = d$$

we obtain from (4.8) that

$\dim R(P_0(\varepsilon)\ldots P_m(\varepsilon))=n-d=\dim N(A_0(\varepsilon))$ for $\varepsilon\in]0,\varepsilon_o]$.

Since $x\in N(A_0(\varepsilon)) \Rightarrow x\in N(A_k(\varepsilon))$ for $k=1,\ldots,m$ and non-zero ε, it follows that. $N(A_0(\varepsilon))=R(P_0(\varepsilon)\ P_1(\varepsilon)\ldots P_m(\varepsilon))$. Therefore, $A_{m+1}(\varepsilon)=0$. Thus (4.7) simplifies to

$$e^{A_0(\varepsilon)t} = \sum_{k=0}^{m} Q_k(\varepsilon)e^{A_k(\varepsilon)\varepsilon^k t} + P_0(\varepsilon)\ldots P_m(\varepsilon) \qquad (4.9)$$

We now prove that $Q_k(\varepsilon), P_k(\varepsilon), A_k(\varepsilon)$ can be replaced by $Q_k=I-P_k$, P_k and A_{k0} respectively. To estimate the difference, we have

$$Q_k(\varepsilon)e^{A_k(\varepsilon)\varepsilon^k t} - Q_k e^{A_{k0}\varepsilon^k t} = \qquad (4.10)$$

$$= \frac{-1}{2\pi i}\int_{\Gamma_k} e^{\lambda\varepsilon^k t}(R(\lambda,A_k(\varepsilon))-R(\lambda,A_{k0}))\,d\lambda$$

where Γ_k is any compact contour enclosing all non-zero eigenvalues of A_{k0}. By the semi-stability of A_{k0}, we can choose Γ_k to be in the left half plane bounded away from the $i\omega$-axis. Thus, for each $\varepsilon>0$, $e^{\lambda\varepsilon^k t}$ is uniformly (in t) bounded on Γ_k. Using this and (4.1) of the lemma, we may conclude that

$$\lim_{\varepsilon\downarrow 0}\ \sup_{t\geq 0}\left\| Q_k(\varepsilon)e^{A_k(\varepsilon)\varepsilon^k t} - Q_k e^{A_{k0}\varepsilon^k t}\right\| = 0.$$

This establishes (3.4). Since, $R(Q_k(\varepsilon))\cong R(A_{k0})$ (3.2) is established. From (3.2) and (3.4), it follows that (3.5), (3.6) and (3.7) are also valid.

Comment (Rows of matrices in Table 1)

From part (iii) of the lemma and the proof above, it is clear that

$$e^{A_0(\varepsilon)t} = \sum_{k=0}^{m} Q_k(\varepsilon)e^{A_k(\varepsilon)\varepsilon^k t} + P_0(\varepsilon)\ldots P_m(\varepsilon) \qquad (4.11)$$

with

$$A_k(\varepsilon) = \sum_{\ell=0}^{\infty} A_{k\ell}\varepsilon^{\ell} \qquad (4.12)$$

where the $A_{k\ell}$ are as given in Table 1. Further, it is clear that for the uniform asymptotic approximation of $e^{A_0(\varepsilon)t}$ as $\varepsilon\downarrow 0$ only the leading terms in (4.12), the $\{A_{k0}\}_{k=1}^{m}$ need be retained.

5. TIGHTNESS OF THE SEMI-STABILITY CONDITION

The requirement of semi-stability for the matrices $A_{00}, A_{10}, \ldots, A_{m0}$ for the system (1.1) to have well defined behavior at time scales $t, t/\varepsilon, \ldots, t/\varepsilon^m$ is tight-counterexamples can be found for the nonexistence of well defined behavior at different time scales if A_{k0} is not semi-stable for some k:

Counterexample 1 (A_{00} does not have SNS)

Let $A_0(\epsilon) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} - \varepsilon\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

Then,

$$e^{A_0(\varepsilon)t} = \begin{bmatrix} \cos\sqrt{\varepsilon}\, t & -\sqrt{\varepsilon}\,\sin\sqrt{\varepsilon}\, t \\ \frac{1}{\varepsilon}\sin\sqrt{\varepsilon}\, t & \cos\sqrt{\varepsilon}\, t \end{bmatrix} e^{-\varepsilon t}$$

Note that $\lim_{\varepsilon\downarrow 0} e^{A_0(\varepsilon)t/\varepsilon}$ does not exist for any t showing lack of well defined behavior at time scale t/ε.

Remark: Recall that A_{k0} is semistable if it has SNS and all non zero eigenvalues have negative real parts. The necessity of the second condition is obvious and we will not furnish a counterexample to illustrate it. We would like to note that the semistability of $A_0(\varepsilon)$ for $\varepsilon\in[0,\varepsilon_o]$ does not imply the semistability of $\{A_{k0}\}_{k=1}^{m}$.

Counterexample 2 (Semistability of $A_0(\varepsilon) \not\Rightarrow$ semistability of A_{k0})

Let

$$A_0(\varepsilon) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & -2 \end{bmatrix} + \varepsilon\begin{bmatrix} 1 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that the eigenvalues of $A_0(\varepsilon)$ are 0, $-2+o(1)$, $-\varepsilon^2+o(\varepsilon^2)$, showing the semi-stability of $A_0(\varepsilon)$. It may further be verified that

$$A_{10} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1/2 & 0 \end{bmatrix}$$

so that A_{10} is nilpotent. In this example $\lim_{\varepsilon\downarrow 0} e^{A_0(\varepsilon)t/\varepsilon^2}$ does not exist showing no well-defined behavior at time scale t/ε^2 in spite of a real eigenvalue of order ε^2.

6. CONCLUSION

Our theorem in section 3 provides a uniform approximation over the entire real line $[0,\infty)$ to the evolution of the system (1.1), thereby extending the results of [1], which are valid only for intervals of the form $[0,T/\varepsilon]$. Furthermore, the hierarchy of models which results from the Corollary is an extension to multiple time scales, of the aggregation results in [6]. The application of these approximations to problems in estimation and control, are currently under study, and will be reported in later publications.

REFERENCES

[1] S.L. Campbell and N.J. Rose, "Singular Perturbations of Autonomous Linear Systems," SIAM J. Math. Anal., Vol. 10, No. 3, p. 542, 1979.

[2] S.L. Campbell, "Singular Perturbations of Autonomous Linear Systems II," J. Diff. Eq., Vol. 29, p. 362, 1978.

[3] M. Coderch and A.S. Willsky, "Hierarchical Aggregation of Finite State Markov Processes," in preparation.

[4] W. Eckhaus, "Asymptotic Analysis of Singular Perturbation," North-Holland, Amsterdam, 1979.

[5] T. Kato, "Perturbation Theory for Linear Operators, Springer Verlag, Berlin, 1966.

[6] P.V. Kokotovic, R.E. O'Malley,Jr. and P. Sannuti, "Singular Perturbations and Order Reduction in Control Theory-An Overview," Automatica, Vol. 12, p.123, 1976.

[7] S.S. Sastry and C.A. Desoer, "Asymptotic Unbounded Root Loci Formulae and Computation," College Eng., U. Berkely, Memo UCB/ERL M81/6 Dec. 1980.

$$
\begin{array}{llll}
A_{00} & A_{01} & \cdots\cdots\cdots & A_{0,\ell} \\
A_{10} & A_{11} & \cdots\cdots & A_{1,\ell-1} \\
\vdots & \vdots & & \\
A_{\ell-1,0} & A_{\ell-1,1} & & \\
A_{\ell,0} & & &
\end{array}
$$

TABLE 1: The array $A_{k,\ell}$ is grown triangularly.